

US Air-Force EOARD Contract F61775-00-WE067

Multiscale Computational Methods in Molecular Simulations

Achi Brandt

Dept. of Appl. Math. & Computer Science
The Weizmann Institute of Science
Rehovot 76100, Israel

Final Scientific Report
covering the period
October 1, 2000 – September 30, 2001

November, 2001

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 29-01-2002		2. REPORT TYPE Final Report		3. DATES COVERED (From – To) 18 September 2000 - 18-Nov-01	
4. TITLE AND SUBTITLE Multiscale Computational Methods in Molecular Simulations			5a. CONTRACT NUMBER F61775-00-WE067		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Professor Achi Brandt			5d. PROJECT NUMBER		
			5d. TASK NUMBER		
			5e. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Weizmann Institute of Science Rehovot Rehovot 76 100 Israel				8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 BOX 14 FPO 09499-0014				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) SPC 00-4067	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This report results from a contract tasking The Weizmann Institute of Science as follows: The contractor will investigate: (1) Adjusting and packaging the software for fast electrostatic summations, so that it can be used by other Air Force projects. (2) The improvement of the simple models used for outlining the coarsening methods to more complex bonded molecular systems. (3) The algorithm for improving the efficiency of Density Functional Theory (DFT) has been demonstrated so far only in one dimension and requires further work for three dimensions, and implementation.					
15. SUBJECT TERMS EOARD, Modelling & Simulation, Materials, Molecular Chemistry					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UL	18, NUMBER OF PAGES 10	19a. NAME OF RESPONSIBLE PERSON Alexander J. Glass, Ph. D.
a. REPORT UNCLAS	b. ABSTRACT UNCLAS	c. THIS PAGE UNCLAS			19b. TELEPHONE NUMBER <i>(Include area code)</i> +44 (0)20 7514 4953

The Contractor, The Weizmann Institute of Science, thereby declares that, to the best of its knowledge and belief, the technical data delivered herewith under Contract No. F61775-00-WE067 is complete, accurate, and complies with all requirements of the contract.

DATE: _____

Name and Title of Authorized Official: _____

I certify that there were no subject inventions to declare as defined in FAR 52.227-13, during the performance of this contract.

DATE: _____

Name and Title of Authorized Official: _____

Abstract

A major bottleneck in material sciences, and in most other branches of applied chemistry, is the inability of current computational methods to simulate large-scale molecular systems at equilibrium or over sufficiently large time intervals. This inability is caused by the product of several complexity factors, associated with the wide range of space and time scales characteristic to such systems. Multiscale computational methods, applied in other fields and tested also for a range of molecular models, have proven able in principle to remove each of these complexity factors. The present research involved further development of these methods, toward realistic molecular systems. This requires a very systematic research and laborious development, climbing level after level, from the microscopic scales to increasingly larger ones, constructing the degrees of freedom, equations and operation rules that efficiently describe each scale of the material. Major advances have been achieved in fast simulation of macromolecules, model multiscaling of fluids, renormalization-multigrid (RMG) methods for dynamic systems, and a new, more general algebraic-multigrid (AMG) solver suitable for implementing the multiscale-eigenbasis (MEB) fast computation of electronic structures.

Scientific Summary

Under the present EOARD grant, aided by other AF funds and by resources at the Weizmann Institute, over the past year we have continued a multiscale program for computational chemistry. Achievements include the following.

1. Fast summation of electrostatic forces

We had previously demonstrated the efficiency of a method suggested in [2] based on a decomposition of the two-particle potential, at each scale, into a local part and a smooth part. An important advantage of this highly-parallelizable approach is that it gives the kind of multiscale force-field description which is needed not only for its fast summation, but also for the efficient multiscale of atomic *motions*. In particular, it is used in investigations reported below (Secs. 2 and 3(B)).

Several important further developments, partly reported in [14] and [13], include: (i) Generalization of the method to fields generated by *dipoles*, in addition to those created by charges. (ii) Substantially higher accuracy for negligible additional CPU time. This has been obtained by introducing enhanced interpolation orders and longer softening distances *only at the coarser levels*, and by correcting for some false *self-interaction*, i.e., the residual interaction of a charge with itself, caused by the multiscale calculations. (iii) For an optional modest increase in CPU time, orders-of-magnitude higher accuracy can be obtained by a new routine that also corrects for errors in the interaction of a particle (or a gridpoint, at coarser levels) with its *closest neighbors*.

2. Multiscale Monte-Carlo methods for macromolecules

In our coarsening method for multiscale simulation of equilibria at a given (e.g., room) temperature, each coarse-level “atom” stands for the average location of several (m , say) next-finer-level atoms. The coarse-level Hamiltonian is developed by extensive, but only local, Monte-Carlo simulations, iteratively adjusting coarse-level Lennard-Jones-type interactions, distribution functions and correlation coefficients so as to fit coarse-level with fine-level simulations. This calculation need be done only once for all similarly-structured molecular neighborhoods. Due to the special choice of internal coordinates, most inter-coordinate coarse-level correlations can be neglected, yielding relatively simple coarse-level Hamiltonians.

Methods for the fast convergence of the iterative adjustments have been developed, including rules for adding new terms to the coarse Hamiltonian and for correcting coefficients of existing terms. They are reported in [1], along with numerical experiments on a united-atom polymer model involving all chemical bonding and Lennard-Jones interactions. A general criterion for selecting coarse-level variables has been developed which is based on the speed of equilibration of “*compatible*” *Monte Carlo* (a MC process at the

fine level restricted to configurations compatible with one given coarse-level configuration). Using this tool it has been determined that the coarsening ratio m at the first level of the united atom model should be 2 or 3. (Indeed, $m > 3$ would not fix enough degrees of freedom to yield unique values to all the dihedral angles, hence would result in slowly equilibrating compatible MC.)

Basic achievement. Numerical experiments with this coarsening ratio and with only one coarse level show that the simulations are already accelerated by a factor larger than 10, while still producing statistics that very accurately match fine-level simulations — for chains small enough to afford simulations at the fine level. This is a basic achievement: it means that the coarse level successfully averages over the local attraction basins of the dihedrals.

Current investigation. The *next* coarsening level (level 2) indicates a similar gain (another factor 10) in efficiency. However, the larger chains that can now be simulated to compare the new level (2) with the previous one (level 1) showed serious inaccuracies, implying the need to add new types of Hamiltonian terms related to *folded* chains. Such terms include non-radial Lennard-Jones (LJ) components as well as correlations between the LJ forces and the internal coordinates. The detailed research of such terms is in progress.

Methods are also being developed for polymer models that involve electrostatic interactions. Based on our approach for fast summation (see above), the *local part* of these interactions participates in the local MC determination of the internal-coordinate terms of the coarse Hamiltonian, while the *smooth part* is computed directly in Cartesian coordinates of the coarse atoms, onto which the electrostatic charges and dipoles are interpolated.

3. Multiscale Monte-Carlo for fluids

Direct fine-scale MC simulations of fluids tend to be extremely inefficient due to the very slow change of various kinds of *clusters* at various scales. The main two kinds of clustering difficulties associated with water and other fluids are *positional* clustering and *electrostatic* ones. We have started out by studying in parallel the following two simple models, in each of which only one of these difficulties is present.

A. Lennard-Jones fluid, featuring positional clustering. While at the finest level we have atoms at arbitrary locations, the coarse levels that we introduce are each defined on a uniform *lattice*. The value defined at each lattice point stands for the average fluid density in a box around that point. The probability distribution of this density depends on neighboring-point densities, as specified in detail by a suitable “*conditional-probability (CP) table*”. Following the Renormalization Multigrid (RMG) methodology developed earlier [9, 10], the CP table for each level is calculated by extensive local simulations at the next finer level. Iterations between the levels is used both for accelerating

the Monte-Carlo simulation on each of them, and for inexpensive calculations of large-scale statistics and macroscopic “equations” (in the form of CP tables). Tests with simple cases are reported in [6], [4, §14.7] and [7], demonstrating accurate predictions of density fluctuations at all scales, with only relatively small grids being simulated at each level, and without MC slowdowns.

The emphasis so far has been on one-dimensional models for which there exist analytical solutions to compare with. This work resulted in clarifying various general algorithmic issues, such as the branching structure of the CP tables, methods for interpolating bin statistics at high particle-per-cell numbers, the employment of only small fine-level simulation domains placed as “windows” in a large physical domain of the next-coarser level, the self-consistent mutual equilibration of all the levels, special procedures at low temperatures [4, §14.7.3], etc.

B. Fixed-location rotating dipoles, to study electrostatic clustering patterns. The chosen model features a large set of electrostatic dipoles, of given strengths and *fixed* locations, rotating in their mutual fields in thermal equilibrium. Clusters of aligned dipoles tend to form, their sizes depending on the given temperature. Again, these clusters are very slow to change in ordinary MC simulations, making large-scale fluctuations extremely slow to average out.

In the multiscale algorithm that we have developed for this model, each coarse level is again defined on a lattice, the mesh size being doubled at each coarsening. With this type of coarsening, the RMG methodology is again applied.

As a first step, we have applied the compatible Monte Carlo (CMC) test mentioned above to choose suitable coarse-level *variables*. This task proved less easy here due to a competition between two conflicting orders: while external or far-field electrostatic interactions tend to *align* neighboring dipoles, their short-range interactions tend to *anti-align* the dipole component perpendicular to the direction of a neighbor. As a result, two types of coarse-level variables proved each inadequate *by itself* (gave unsatisfactory CMC convergence speed). We then tried a combination of the two: the coarse level includes both “*injections*” (i.e., one part of the coarse variables is simply a *subset* of the fine-level variables) and “*super-dipoles*” (interpolated from the next-fine-level dipoles). This yielded good coarsening, and it also taught us that our *CMC test should generally be extended*: At a low temperature, fast equilibration may not be obtained with simple CMC, but it should be obtained by a CMC process started at a higher temperature and then lowered (in few steps) to the given low temperature. We call this more general process *quick-annealing CMC* (qaCMC). We expect it to be an important coarse-to-fine interpolation process in many problems.

4. Dynamic RMG

The methods described above are restricted to systems *in equilibrium*. To calculate

time-dependent processes, an extension of the Renormalization Multigrid (RMG) methodology to dynamic systems should be developed. In the same way that our algorithmic concepts for equilibrium systems had been first developed for simple lattice models, we have now embarked on studying RMG techniques for simple *driven* (i.e., non-equilibrium) diffusive lattice systems (see, e.g., [15]) in one dimension. The systems consist of particles at lattice points; at each time step each particle hops to a neighboring point with a certain probability that depends, e.g., on its distance from other particles.

We have demonstrated that the RMG techniques can be extended to describe such systems at increasingly coarser levels with correspondingly larger time steps. At each coarse level, the dependence of fluxes (net particle rates of flow between adjacent coarse-level boxes) on neighboring fluxes and densities is derived by Monte-Carlo simulations at the next finer level and stored in “Conditional Probability (CP) tables”. With such a table one can then simulate the coarser level and thereby derive CP tables for the next, still coarser level; etc. The tests show that with branching tables similar to those in [9, 10], coarse levels can very inexpensively and very accurately reproduce large-scale statistics.

This work in progress opens the way to efficient coarsening not only for systems with stochastic dynamics (such as molecular Langevin dynamics at a given finite temperature) but also for systems with deterministic dynamics that at much larger scales become practically stochastic.

A particular type of this RMG technique can be used to *derive equilibrium “equations”* (in the form of a *fine-scale* CP table) from the dynamic model, by coarsening (doubling) at each renormalization transformation only the time step, while leaving the spatial scale (the size of the boxes) fixed. At the limit of such transformation, steady-state CP tables should emerge.

5. Electronic structure calculation: algebraic multigrid

A new method had previously been developed for calculating a large number N of eigenfunctions of a differential operator to accuracy ε in just $O(N(\log N)(\log \frac{1}{\varepsilon}))$ operations, instead of $O(N^3)$ needed before (see Sec. 9.2 in [3]). In collaboration with Ph.D. student Oren Livne, this performance had been demonstrated on simple model problems (1D Schrödinger equations), as reported in [11] and [12]. However, that feasibility demonstration had been based on formulations unique to the one-dimensional case. In particular, in 1D it is possible to avoid solving highly indefinite boundary-value problems, thus skipping the more involved mechanism required for coarsening such problems (see [8]). The extension to higher dimensions (discussed in [12]) is far from trivial, and intimately related to the extension of the wave/ray multigrid methods to variable coefficients (see [4, §7.1]) and to general matrices (see [4, §17.2.2]). This requires, first of all, the further development of the new concept of “*bootstrap algebraic multigrid*” (*BAMG*), introduced in [4, §17.2].

Like previous algebraic multigrid (AMG) approaches, this one involves methods for deriving coarser level equations for a given (fine-level) system of linear equations by using the Galerkin principle based on simple inter-level transfers. While previous AMG schemes were limited to special types of matrices, the BAMG algorithm introduces systematic principles for constructing inter-grid transfers which are applicable to any system of local equations. The inter-grid transfers are based on near-eigenfunctions, hence their construction is closely related to the “multiscale eigenbasis” (MEB) needed for the $O(N \log N)$ calculation of N eigenfunctions, e.g., for solving the equations of the density-functional theory (DFT) of electronic structures.

Thus, the effort at this stage, in collaboration with Dr. Oren Livne, is focused on the detailed development of the BAMG approach, which has very wide applications, including fast solvers for electronic structures.

References

- [1] Bai, D. and Brandt, A., Multiscale computation of polymer models, in [5], pp. 250–266.
- [2] Brandt, A., Multilevel computations of integral transforms and particle interactions with oscillatory kernels, *Comp. Phys. Comm.* **65** (1991) 24–38.
- [3] Brandt, A., Multiscale scientific computation: six year summary, Gauss Center Report WI/GC-12, May 1999. Also Appeared in MGNET.
- [4] Brandt, A., Multiscale scientific computation: review 2001, in *Multiscale and Multiresolution Methods* (T. Barth, R. Haimes and T. Chan, eds.), Springer Verlag, Heidelberg, 2001 (*Proc. Yosemite Educational Symp.*, October 2000).
- [5] Brandt, A., Bernholc, J. and Binder, K. (Eds.), *Multiscale Computational Methods in Chemistry and Physics*, NATO Science Series: Computer and System Sciences, Vol. 177, IOS Press, Amsterdam, 2001.
- [6] Brandt, A. and Iliyn, V., Multilevel approach in statistical physics of liquids, in [5], pp. 187–197.
- [7] Brandt, A. and Iliyn, V., Multilevel Monte Carlo methods for studying large-scale phenomena in fluids, *J. Molecular Liquids*, Elsevier Science, to appear in 2002.
- [8] Brandt, A. and Livshits, I., Wave-ray multigrid method for standing wave equations, *Electronic Trans. Num. An.* **6** (1997) 162–181.
- [9] Brandt, A. and Ron, D., Renormalization multigrid (RMG): Statistically optimal renormalization group flow and coarse-to-fine Monte Carlo acceleration, Gauss Center Report WI/GC-11, June 1999. *J. Stat. Phys.* **102** (2001) 231–257.
- [10] Brandt, A. and Ron, D., Renormalization multigrid (RMG): coarse-to-fine Monte Carlo acceleration and optimal derivation of macroscopic descriptions, in [5], pp. 163–186.
- [11] Livne, O. E., *Multiscale Eigenbasis Algorithms*, Ph.D. Thesis, Weizmann Institute of Science, Rehovot, 2000.
- [12] Livne, O. and Brandt, A., $O(N \log N)$ multilevel calculation of N eigenfunctions, in [5], pp. 112–136.

- [13] Sandak, B., Efficient computational algorithms for fast electrostatic and molecular docking, *Proc. 3rd Int. Workshop on Methods for Macromolecular Modeling* (New York, October 12–14, 2000), *Springer Verlag Lecture Notes in Computational Science and Engineering*.
- [14] Sandak, B. and Brandt, A., Multiscale fast summation of long range charge and dipolar interactions, in [5], pp. 6–31. Also: *J. Comp. Chem.* **22** (2001) 717–731.
- [15] Schmittmann, B. and Zia, R. K. P., *Statistical Mechanics of Driven Diffusive Systems*, Academic Press, 1995.